

CLAIMS

What is claimed is:

- 1 1. A circuit arrangement for searching a parent code sequence for a target code
2 sequence, comprising:
3 a shift register arrangement having a plurality of stages, wherein each stage stores
4 a code of a subset of codes of the parent code sequence, and the shift register arrangement
5 is adapted to periodically shift the subset of codes to form a new subset of codes with
6 another code from the parent code sequence in a leading stage;
7 a matching circuit coupled to the shift register arrangement, the matching circuit
8 adapted to ascertain code position matches between the subset of codes in the stages of the
9 shift register arrangement and codes in corresponding code positions of the target code
10 sequence, and provide a programmed binary value for each code position match; and
11 a pipelined adder arrangement coupled to the matching circuit, the adder
12 arrangement adapted to sum the binary values for code position matches for each
13 respective subset of codes.
- 1 2. The circuit arrangement of claim 1, wherein each stage of the shift register
2 arrangement is adapted for storage of a code of character data.
- 1 3. The circuit arrangement of claim 1, wherein each stage of the shift register
2 arrangement is adapted for storage of a code of a plurality of character data.
- 1 4. The circuit arrangement of claim 1, wherein the pipelined adder arrangement is a
2 pipelined adder tree.
- 1 5. The circuit arrangement of claim 1, wherein the pipelined adder arrangement
2 includes at least one stage of pipelined carry-save adders coupled to at least one stage of
3 pipelined carry-propagate adders.
- 1 6. The circuit arrangement of claim 5, wherein the at least one stage of pipelined
2 carry-save adders are adapted to provide a plurality of binary vectors responsive to the

3 quantity of code position matches, and the at least one stage of pipelined carry-propagate
4 adders are adapted to add the plurality of binary vectors.

1 7. The circuit arrangement of claim 1, further comprising a pipelined summing circuit
2 coupled to the pipelined adder arrangement and adapted to determine a moving sum of
3 code position matches for a plurality of subsets of codes.

1 8. The circuit arrangement of claim 7, wherein the plurality of subsets of codes
2 includes at least a most recent subset of codes and a next most recent subset of codes.

1 9. The circuit arrangement of claim 7, wherein the plurality of subsets of codes
2 includes a first subset of codes and a prior subset of codes, wherein an intervening subset
3 of codes is processed between the first subset of codes and the prior subset of codes.

1 10. The circuit arrangement of claim 1, wherein each subset of codes includes n
2 contiguous codes from the parent code sequence.

1 11. The circuit arrangement of claim 1, wherein the matching circuit includes a
2 plurality of programmable lookup tables, each lookup table having an input terminal
3 coupled to an output terminal of a corresponding stage of the shift register arrangement
4 and configured to provide a programmed value responsive to an input code value.

1 12. A method for searching a parent code sequence for a target code sequence,
2 comprising:
3 shifting the parent code sequence through a shift register arrangement having a
4 plurality of stages, wherein the shift register arrangement stores a subset of codes of the
5 parent code sequence and each stage stores a code of the subset of codes, and each shift of
6 the subset of codes forms a new subset of codes with another code from the parent code
7 sequence in a leading stage;
8 determining in parallel whether the codes in the stages of the shift register
9 arrangement are equal to codes of the target code sequence in corresponding code

10 positions, and generating in parallel signals of a programmed binary value for each
 11 equality of a subset code and a target code; and
 12 summing the signals of the programmed binary value in a pipelined adder that
 13 generates a sum corresponding to each shift of the shift register arrangement.

1 13. The method of claim 12, further comprising:
 2 determining, for each respective subset of codes, a probability of being the target
 3 code sequence as the sum of the binary values for code position matches for the respective
 4 subset of codes divided a total quantity of code positions in the target code sequence; and
 5 associating the probability for each respective subset of codes with a unique
 6 identifier representative of a location within the parent code sequence at which the
 7 respective subset of codes exists.

1 14. The method of claim 12, wherein the parent code sequence represents a genome.

1 15. The method of claim 14, wherein each code of the parent code sequence is
 2 representative of a nucleotide type.

1 16. The method of claim 15, wherein the nucleotide type is selected from the group
 2 consisting of: adenine, thymine, guanine, and cytosine.

1 17. The method of claim 14, wherein the genome is a human genome.

1 18. The method of claim 12, further comprising:
 2 configuring a plurality of lookup tables to generate respective signals of the
 3 programmed binary value when addressed by codes equal to codes of the target code
 4 sequence; and
 5 addressing the lookup tables with the codes of the subset of codes.

1 19. The method of claim 12, further comprising generating a moving sum, for n
 2 subsets of codes, of sums of the signals of the selected binary value.

1 20. The method of claim 19, wherein the n subsets of codes includes at least a most
2 recent subset of codes and a next most recent subset of codes.

1 21. The method of claim 19, wherein the n subsets of codes includes a first subset of
2 codes and a prior subset of codes, wherein an intervening subset of codes is processed
3 between the first subset of codes and the prior subset of codes.

1 22. The method of claim 12, wherein each subset of codes includes m contiguous
2 codes from the parent code sequence.

1 23. An apparatus for searching a parent code sequence for a target code sequence, each
2 code in the parent code sequence having a parent-relative position, comprising:
3 means for periodically selecting subsets of codes of the parent code sequence, each
4 code in the subset having a relative subset-code position defined by the parent-relative
5 position, and each subset of codes differing from other subsets by parent-relative positions
6 of the codes in the subset;
7 means for determining in parallel whether each code at a subset-code position in a
8 subset of codes is equal to a code of the target code sequence in a corresponding target-
9 code position, and generating in parallel signals of a selected binary value for each
10 equality of a subset code and the target code; and
11 means for summing the signals of the selected binary value.